RESULT   3
AE000166/c
LOCUS   AE000166   15505 bp   DNA   BCT   01-DEC-2000
DEFINITION   Escherichia coli K12 MG1655 section 56 of 400 of the complete
             genome.
ACCESSION   AE000166 U00096
VERSION   AE000166.1 GI:1786819
KEYWORDS   .
SOURCE   Escherichia coli K12.
  ORGANISM   Escherichia coli K12
             Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
             Escherichia.
REFERENCE   1 (bases 1 to 15505)
  AUTHORS   Blattner,F.R., Plunkett,G. III, Bloch,C.A., Perna,N.T., Burland,V.,
             Riley,M., Collado-Vides,J., Glasner,J.D., Rode,C.K., Mayhew,G.F.,
             Gregor,J., Davis,N.W., Kirkpatrick,H.A., Goeden,M.A., Rose,D.J.,
             Mau,B. and Shao,Y.
  TITLE   The complete genome sequence of Escherichia coli K-12
  JOURNAL   Science 277 (5331), 1453-1474 (1997)
  MEDLINE   97426617
   PUBMED   9278503
REFERENCE   2 (bases 1 to 15505)
  AUTHORS   Blattner,F.R.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics,
             University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
             Email: ecoli@genetics.wisc.edu Phone: 608-262-2534 Fax:
             608-263-7459
REFERENCE   3 (bases 1 to 15505)
  AUTHORS   Blattner,F.R.
  TITLE   Direct Submission
  JOURNAL   Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics,
             University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
             Email: ecoli@genetics.wisc.edu Phone: 608-262-2534 Fax:
             608-263-7459
REFERENCE   4 (bases 1 to 15505)
  AUTHORS   Plunkett,G. III.
  TITLE   Direct Submission
  JOURNAL   Submitted (13-OCT-1998) Laboratory of Genetics, University of
             Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
COMMENT   This sequence was determined by the E. coli Genome Project at the
             University of Wisconsin-Madison (Frederick R. Blattner, director).
             Supported by NIH grants HG00301 and HG01428 (from the Human Genome
             Project and NCHGR). The entire sequence was independently
             determined from E. coli K12 strain MG1655. Predicted open reading
             frames were determined using GeneMark software, kindly supplied by
             Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA,
             30332 [e-mail: mark@amber.gatech.edu]. Open reading frames that
             have been correlated with genetic loci are being annotated with CG
             Site Nos., unique ID nos. for the genes in the E. coli Genetic
             Stock Center (CGSC) database at Yale University, kindly supplied by
             Mary Berlyn. A public version of the database is accessible
             (http://cgsc.biology.yale.edu). Annotation of the genome is an
             ongoing task whose goal is to make the genome sequence more useful
             by correlating it with other data. Comments to the authors are
             appreciated. Updated information will be available at the E. coli
             Genome Project's World Wide Web site
             (http://www.genetics.wisc.edu). *** The E. coli K12 sequence and
             its annotations are periodically updated; this is version M54. No
             sequence changes. Annotation updates: updated gene identifications
             and products; all new functional assignments courtesy of Monica
             Riley; added promoters, protein binding sites, and repeated
             sequences described in reference 1. The unique numeric identifiers
             beginning with a lowercase 'b' assigned to each gene (protein- or
             RNA-encoding) are now designated as gene synonyms instead of
             labels. This should allow them to be searched for in Entrez as gene
             names.
FEATURES   Location/Qualifiers
     source   1. .15505

```
                        /organism="Escherichia coli K12"
                        /strain="K12"
                        /sub_strain="MG1655"
                        /db_xref="taxon:83333"
     gene               complement(104. .1006)
                        /gene="ybdO"
                        /note="b0603"
     CDS                complement(104. .1006)
                        /gene="ybdO"
                        /function="putative regulator; Not classified"
                        /note="f300; This 300 aa ORF is 29 pct identical (5 gaps)
                        to 295 residues of an approx. 320 aa protein YBEF_ECOLI
                        SW: P30979"
                        /codon_start=1
                        /transl_table=11
                        /product="putative transcriptional regulator LYSR-type"
                        /protein_id="AAC73704.1"
                        /db_xref="GI:1786820"
                        /translation="MANLYDLKKFDLNLLVIFECIYQHLSISKAAESLYITPSAVSQS
                        LQRLRAQFNDPLFIRSGKGIAPTTTGLNLHHHLEKNLRGLEQTINIVNKSELKKNFII
                        YGPQLISCSNNSMLIRCLRQDSSVEIECHDILMSAENAEELLVHRKADLVITQMPVIS
                        RSVICMPLHTIRNTLICSNRHPRITDNSTYEQIMAEEFTQLISKSAGVDDIQMEIDER
                        FMNRKISFRGSSLLTIINSIAVTDLLGIVPYELYNSYRDFLNLKEIKLEHPLPSIKLY
                        ISYNKSSLNNLVFSRFIDRLNESF"
     promoter           complement(1007. .1034)
                        /note="factor Sigma70; predicted +1 start at 636835"
     promoter           complement(1018. .1046)
                        /note="factor Sigma70; predicted +1 start at 636846"
     promoter           complement(1045. .1071)
                        /note="factor Sigma70; predicted +1 start at 636873"
     promoter           complement(1081. .1110)
                        /note="factor Sigma70; predicted +1 start at 636909"
     promoter           complement(1166. .1196)
                        /note="factor Sigma70; predicted +1 start at 636994"
     gene               complement(1215. .2021)
                        /gene="dsbG"
                        /note="b0604"
     CDS                complement(1215. .2021)
                        /gene="dsbG"
                        /function="enzyme; Central intermediary metabolism: Pool,
                        multipurpose conversions"
                        /note="f268; This 268 aa ORF is 30 pct identical (31 gaps)
                        to 220 residues of an approx. 240 aa protein DSBC_ECOLI
                        SW: P21892"
                        /codon_start=1
                        /transl_table=11
                        /product="thiol:disulfide interchange protein"
                        /protein_id="AAC73705.1"
                        /db_xref="GI:1786821"
                        /translation="MTVIGYAFYSTFALTEKDKLMLKKILLLALLPAIAFAEELPAPV
                        KAIEKQGITIIKTFDAPGGMKGYLGKYQDMGVTIYLTPDGKHAISGYMYNEKGENLSN
                        TLIEKEIYAPAGREMWQRMEQSHWLLDGKKDAPVIVYVFADPFCPYCKQFWQQARPWV
                        DSGKVQLRTLLVGVIKPESPATAAAILASKDPAKTWQQYEASGGKLKLNVPANVSTEQ
                        MKVLSDNEKLMDDLGANVTPAIYYMSKENTLQQAVGLPDQKTLNIIMGNK"
     promoter           complement(2127. .2155)
                        /note="factor Sigma70; predicted +1 start at 637955"
     promoter           complement(2176. .2206)
                        /note="factor Sigma70; predicted +1 start at 638004"
     promoter           2232. .2261
                        /note="factor Sigma70; predicted +1 start at 638103"
     protein_bind       2247. .2260
                        /note="central position to predicted promoter: -13"
                        /bound_moiety="OxyR predicted site"
     gene               2333. .2896
                        /gene="ahpC"
                        /note="b0605"
     CDS                2333. .2896
                        /gene="ahpC"
                        /EC_number="1.6.4.-"
                        /function="enzyme; Detoxification"
                        /note="o187; 100 pct identical to AHPC_ECOLI SW: P26427"
```

```
                          /codon_start=1
                          /transl_table=11
                          /product="alkyl hydroperoxide reductase, C22 subunit;
                          detoxification of hydroperoxides"
                          /protein_id="AAC73706.1"
                          /db_xref="GI:1786822"
                          /translation="MSLINTKIKPFKNQAFKNGEFIEITEKDTEGRWSVFFFYPADFT
                          FVCPTELGDVADHYEELQKLGVDVYAVSTDTHFTHKAWHSSSETIAKIKYAMIGDPTG
                          ALTRNFDNMREDEGLADRATFVVDPQGIIQAIEVTAEGIGRDASDLLRKIKAAQYVAS
                          HPGEVCPAKWKEGEATLAPSLDLVGKI"
     repeat_region        2896. .3022
                          /note="IRU"
     gene                 3111. .4706
                          /gene="ahpF"
                          /note="b0606"
     CDS                  3111. .4706
                          /gene="ahpF"
                          /function="enzyme; Detoxification"
                          /note="o531; 100 pct identical to fragment AHPF_ECOLI
                          SW:P35340 (162 aa) but contains additional N-ter and C ter
                          residues"
                          /codon_start=1
                          /transl_table=11
                          /product="alkyl hydroperoxide reductase, F52a subunit;
                          detoxification of hydroperoxides"
                          /protein_id="AAC73707.1"
                          /db_xref="GI:1786823"
                          /translation="MMMFKAQEINMLDTNMKTQLKAYLEKLTKPVELIATLDDSAKSA
                          EIKELLAEIAELSDKVTFKEDNSLPVRKPSFLITNPGSNQGPRFAGSPLGHEFTSLVL
                          ALLWTGGHPSKEAQSLLEQIRHIDGDFEFETYYSLSCHNCPDVVQALNLMSVLNPRIK
                          HTAIDGGTFQNEITDRNVMGVPAVFVNGKEFGQGRMTLTEIVAKIDTGAEKRAAEELN
                          KRDAYDVLIVGSGPAGAAAAIYSARKGIRTGLMGERFGGQILDTVDIENYISVPKTEG
                          QKLAGALKVHVDEYDVDVIDSQSASKLIPAAVEGGLHQIETASGAVLKARSIIVATGA
                          KWRNMNVPGEDQYRTKGVTYCPHCDGPLFKGKRVAVIGGGNSGVEAAIDLAGIVEHVT
                          LLEFAPEMKADQVLQDKLRSLKNVDIILNAQTTEVKGDGSKVVGLEYRDRVSGDIHNI
                          ELAGIFVQIGLLPNTNWLEGAVERNRMGEIIIDAKCETNVKGVFAAGDCTTVPYKQII
                          IATGEGAKASLSAFDYLIRTKTA"
     gene                 complement(4827. .5255)
                          /gene="ybdQ"
                          /note="b0607"
     CDS                  complement(4827. .5255)
                          /gene="ybdQ"
                          /function="orf; Unknown"
                          /note="f142; This 142 aa ORF is 30 pct identical (4 gaps)
                          to 126 residues of an approx. 152 aa protein YFMU_COXBU
                          SW: P45680"
                          /codon_start=1
                          /transl_table=11
                          /product="orf, hypothetical protein"
                          /protein_id="AAC73708.1"
                          /db_xref="GI:1786824"
                          /translation="MYKTIIMPVDVFEMELSDKAVRHAEFLAQDDGVIHLLHVLPGSA
                          SLSLHRFAADVRRFEEHLQHEAQERLQTMVSHFTIDPSRIKQHVRFGSVRDEVNELAE
                          ELGADVVVIGSRNPSISTHLLGSNASSVIRHANLPVLVVR"
     promoter             complement(5274. .5303)
                          /note="factor Sigma70; predicted +1 start at 641102"
     promoter             5275. .5303
                          /note="factor Sigma70; predicted +1 start at 641145"
     gene                 5476. .6714
                          /gene="ybdR"
                          /note="b0608"
     CDS                  5476. .6714
```

```
     Query Match              100.0%;  Score 5484;  DB 1;  Length 15505;
     Best Local Similarity    100.0%;  Pred. No. 0;
     Matches 5484;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 gaacaacattcgttgcaaatcgataacaacatgcaccttcaggatactatttattatgtt 60
            ·|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    15299 GAACAACATTCGTTGCAAATCGATAACAACATGCACCTTCAGGATACTATTTATTATGTT 15240
```

```
Qy    ' ' 61  cggcaatgatatttttcacccgcgtaaaacgttcagaaaataaaaaaatggcggaaatcgc  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    15239  CGGCAATGATATTTTCACCCGCGTAAAACGTTCAGAAAATAAAAAAATGGCGGAAATCGC  15180

Qy     121  ccaattcctgcatgaaaatgatttgagcgttgacaccacagtcgaagtatttattaccgt  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    15179  CCAATTCCTGCATGAAAATGATTTGAGCGTTGACACCACAGTCGAAGTATTTATTACCGT  15120

Qy     181  aacccgcgatgaaaagcttatcgcgtgcggtggaattgccggaaatattattaaatgcgt  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    15119  AACCCGCGATGAAAAGCTTATCGCGTGCGGTGGAATTGCCGGAAATATTATTAAATGCGT  15060

Qy     241  tgctatcagtgaatccgtccgcggtgaaggactggcgctgacattagccactgaattgat  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    15059  TGCTATCAGTGAATCCGTCCGCGGTGAAGGACTGGCGCTGACATTAGCCACTGAATTGAT  15000

Qy     301  aaacctcgcctatgagcggcacagcacgcatctgtttatttataccaaaaccgaatacga  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    14999  AAACCTCGCCTATGAGCGGCACAGCACGCATCTGTTTATTTATACCAAAACCGAATACGA  14940

Qy     361  ggcgctgttccgccagtgcggttttttccacgctgaccagcgtacccggcgtgatggtgct  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    14939  GGCGCTGTTCCGCCAGTGCGGTTTTTTCCACGCTGACCAGCGTACCCGGCGTGATGGTGCT  14880

Qy     421  gatggaaaacagcgccacgcgactgaaacgctatgccgaatcgctgaaaaaatttcgtca  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    14879  GATGGAAAACAGCGCCACGCGACTGAAACGCTATGCCGAATCGCTGAAAAAATTTCGTCA  14820

Qy     481  tccagggaacaagattggctgcattgtgatgaacgccaatccctttacgaatggtcaccg  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    14819  TCCAGGGAACAAGATTGGCTGCATTGTGATGAACGCCAATCCCTTTACGAATGGTCACCG  14760

Qy     541  ttatctgattcaacaggctgcggcacagtgcgactggttgcatctgtttttagtcaaaga  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    14759  TTATCTGATTCAACAGGCTGCGGCACAGTGCGACTGGTTGCATCTGTTTTTAGTCAAAGA  14700

Qy     601  agattcttcacgcttcccctatgaagaccggctggatttggtgttaaaaggcaccgccga  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    14699  AGATTCTTCACGCTTCCCCTATGAAGACCGGCTGGATTTGGTGTTAAAAGGCACCGCCGA  14640

Qy     661  tattccacgcctgactgtgcatcgtggctccgaatacatcatctcccgcgctacgttccc  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    14639  TATTCCACGCCTGACTGTGCATCGTGGCTCCGAATACATCATCTCCCGCGCTACGTTCCC  14580

Qy     721  ttgctacttcattaaagaacagagcgtcattaaccattgttacaccgaaattgatctgaa  780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    14579  TTGCTACTTCATTAAAGAACAGAGCGTCATTAACCATTGTTACACCGAAATTGATCTGAA  14520

Qy     781  gattttccgtcagtacctcgctcccgcgctgggtgtaactcaccgctttgtcggtactga  840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    14519  GATTTTCCGTCAGTACCTCGCTCCCGCGCTGGGTGTAACTCACCGCTTTGTCGGTACTGA  14460

Qy     841  acccttttgtcgcgttaccgcccagtacaaccaggatatgcgctactggctggaaacgcc  900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    14459  ACCCTTTTGTCGCGTTACCGCCCAGTACAACCAGGATATGCGCTACTGGCTGGAAACGCC  14400

Qy     901  gactatctccgcaccgcccatcgaactggttgaaattgagcggctgcgttaccaggagat  960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    14399  GACTATCTCCGCACCGCCCATCGAACTGGTTGAAATTGAGCGGCTGCGTTACCAGGAGAT  14340

Qy     961  gccgatatccgcttcccgggtacgtcaactgctggcgaaaaacgatctcacggctatcgc  1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    14339  GCCGATATCCGCTTCCCGGGTACGTCAACTGCTGGCGAAAAACGATCTCACGGCTATCGC  14280

Qy    1021  gccgctggtccctgcagtcacgctgcattatttgcagaacctgcttgagcactcccgcca  1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    14279  GCCGCTGGTCCCTGCAGTCACGCTGCATTATTTGCAGAACCTGCTTGAGCACTCCCGCCA  14220

Qy    1081  ggacgcggcagctcgtcaaaagacccccgcatgagaaacaggtgaaaaatgaaaataaac  1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    14219  GGACGCGGCAGCTCGTCAAAAGACCCCCGCATGAGAAACAGGTGAAAAATGAAAATAAAC  14160
```

```
Qy   1141  cagcccgccgttgcaggcacccttgagtctggggatgtgatgatacgcatcgccccactc 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  14159  CAGCCCGCCGTTGCAGGCACCCTTGAGTCTGGGGATGTGATGATACGCATCGCCCCACTC 14100

Qy   1201  gatacgcaggatatcgacctgcaaatcaatagcagcgttgagaaacagtttggcgatgca 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  14099  GATACGCAGGATATCGACCTGCAAATCAATAGCAGCGTTGAGAAACAGTTTGGCGATGCA 14040

Qy   1261  attcgcaccaccattctggacgttctcgcccgctacaacgtgcgcggcgtacagctgaat 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  14039  ATTCGCACCACCATTCTGGACGTTCTCGCCCGCTACAACGTGCGCGGCGTACAGCTGAAT 13980

Qy   1321  gtcgatgacaaaggcgcactggactgcattttacgtgcacgactggaagccctgctggca 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  13979  GTCGATGACAAAGGCGCACTGGACTGCATTTTACGTGCACGACTGGAAGCCCTGCTGGCA 13920

Qy   1381  cgcgccagcggtatcccggctctgccatgggaggattgccaatgatttccgcttcgctgc 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  13919  CGCGCCAGCGGTATCCCGGCTCTGCCATGGGAGGATTGCCAATGATTTCCGCTTCGCTGC 13860

Qy   1441  aacaacgtaaaactcgcacccgccgcagcatgttgtttgtgcctggtgccaatgccgcga 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  13859  AACAACGTAAAACTCGCACCCGCCGCAGCATGTTGTTTGTGCCTGGTGCCAATGCCGCGA 13800

Qy   1501  tggtcagcaactccttcatctacccggctgatgccctgatgtttgacctcgaagactccg 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  13799  TGGTCAGCAACTCCTTCATCTACCCGGCTGATGCCCTGATGTTTGACCTCGAAGACTCCG 13740

Qy   1561  tagcattgcgtgaaaaagacaccgcccgccgcatggtttaccacgcgctgcaacatccgc 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  13739  TAGCATTGCGTGAAAAAGACACCGCCCGCCGCATGGTTTACCACGCGCTGCAACATCCGC 13680

Qy   1621  tgtatcgcgatattgaaaccattgtgcgtgtcaacgcgctggattccgaatggggtgtta 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  13679  TGTATCGCGATATTGAAACCATTGTGCGTGTCAACGCGCTGGATTCCGAATGGGGTGTTA 13620

Qy   1681  acgacctggaagccgtcgttcgcggtggtgcggacgttgtgcgtctgccgaaaaccgata 1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  13619  ACGACCTGGAAGCCGTCGTTCGCGGTGGTGCGGACGTTGTGCGTCTGCCGAAAACCGATA 13560

Qy   1741  ccgctcaggatgttctggatattgaaaaagagatcctgcgtatcgaaaaagcctgtggtc 1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  13559  CCGCTCAGGATGTTCTGGATATTGAAAAAGAGATCCTGCGTATCGAAAAAGCCTGTGGTC 13500

Qy   1801  gtgaacccggcagcaccggcctgctggcggcgattgaatctccgctggggattacccgcg 1860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  13499  GTGAACCCGGCAGCACCGGCCTGCTGGCGGCGATTGAATCTCCGCTGGGGATTACCCGCG 13440

Qy   1861  cagtggaaatcgctcacgcttccgagcgtttgatcggtatcgccctcggtgcagaagact 1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  13439  CAGTGGAAATCGCTCACGCTTCCGAGCGTTTGATCGGTATCGCCCTCGGTGCAGAAGACT 13380

Qy   1921  atgtgcgcaacctgcgtacagaacgctccccggaaggaactgaactgctgttcgcacgct 1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  13379  ATGTGCGCAACCTGCGTACAGAACGCTCCCCGGAAGGAACTGAACTGCTGTTCGCACGCT 13320

Qy   1981  gttccattttgcaggccgcgcgctctgcgggtattcaggcgttcgataccgtctattccg 2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  13319  GTTCCATTTTGCAGGCCGCGCGCTCTGCGGGTATTCAGGCGTTCGATACCGTCTATTCCG 13260

Qy   2041  acgctaacaacgaagccggatttctgcaagaagccgcccacatcaaacagctgggctttg 2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  13259  ACGCTAACAACGAAGCCGGATTTCTGCAAGAAGCCGCCCACATCAAACAGCTGGGCTTTG 13200

Qy   2101  acggcaaatcgctgatcaacccgcgtcagattgatctgctgcacaacctctacgcaccga 2160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  13199  ACGGCAAATCGCTGATCAACCCGCGTCAGATTGATCTGCTGCACAACCTCTACGCACCGA 13140

Qy   2161  cccagaaagaagtggatcacgcccgccgcgtcgtagaagccgctgaagccgccgctcgcg 2220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  13139  CCCAGAAAGAAGTGGATCACGCCCGCCGCGTCGTAGAAGCCGCTGAAGCCGCCGCTCGCG 13080
```

```
Qy    2221  aaggcctcggcgtggtttccctgaacggcaagatggtggacggtccggttatcgatcgcg  2280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13079  AAGGCCTCGGCGTGGTTTCCCTGAACGGCAAGATGGTGGACGGTCCGGTTATCGATCGCG  13020

Qy    2281  cccgtctggtgctctcccgtgcagaactttccggcatccgcgaagaataaggcaatcaaa  2340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13019  CCCGTCTGGTGCTCTCCCGTGCAGAACTTTCCGGCATCCGCGAAGAATAAGGCAATCAAA  12960

Qy    2341  atgacgcagaaaattgaacaatctcaacgacaagaacgggtagcggcctggaatcgtcgc  2400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12959  ATGACGCAGAAAATTGAACAATCTCAACGACAAGAACGGGTAGCGGCCTGGAATCGTCGC  12900

Qy    2401  gctgaatgcgatcttgccgcttttccagaactcgccaaagcaaacctaccaggctgaaaaa  2460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12899  GCTGAATGCGATCTTGCCGCTTTCCAGAACTCGCCAAAGCAAACCTACCAGGCTGAAAAA  12840

Qy    2461  gcgcgcgatcgcaaactgtgcgccaacctggaagaagcgattcgtcgctctggtttacag  2520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12839  GCGCGCGATCGCAAACTGTGCGCCAACCTGGAAGAAGCGATTCGTCGCTCTGGTTTACAG  12780

Qy    2521  gacggcatgacggtttccttccatcacgctttccgtggcggtgacctgaccgtcaatatg  2580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12779  GACGGCATGACGGTTTCCTTCCATCACGCTTTCCGTGGCGGTGACCTGACCGTCAATATG  12720

Qy    2581  gtgatggacgtcatcgcgaagatgggctttaaaaacctgaccctggcgtccagctccctg  2640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12719  GTGATGGACGTCATCGCGAAGATGGGCTTTAAAAACCTGACCCTGGCGTCCAGCTCCCTG  12660

Qy    2641  agtgattgccatgcgccgctggtagaacacattcgccagggcgtggttacccgcatttat  2700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12659  AGTGATTGCCATGCGCCGCTGGTAGAACACATTCGCCAGGGCGTGGTTACCCGCATTTAT  12600

Qy    2701  acctccggcctgcgtggtccactggcggaagagatctcccgtggtctgctggcagaaccg  2760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12599  ACCTCCGGCCTGCGTGGTCCACTGGCGGAAGAGATCTCCCGTGGTCTGCTGGCAGAACCG  12540

Qy    2761  gtgcagatccactctcacggcggtcgtgtgcatctggtacagagcggcgaactgaatatc  2820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12539  GTGCAGATCCACTCTCACGGCGGTCGTGTGCATCTGGTACAGAGCGGCGAACTGAATATC  12480

Qy    2821  gacgtggctttcctcggcgtcccgtcctgtgatgaattcggtaatgccaacggctacacc  2880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12479  GACGTGGCTTTCCTCGGCGTCCCGTCCTGTGATGAATTCGGTAATGCCAACGGCTACACC  12420

Qy    2881  ggtaaagcctgctgcggctccctcggctatgcaatagttgatgccgacaacgcaaaacag  2940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12419  GGTAAAGCCTGCTGCGGCTCCCTCGGCTATGCAATAGTTGATGCCGACAACGCAAAACAG  12360

Qy    2941  gtcgtgatgcttaccgaagaactgctgccttatccgcataatccggcaagcattgagcaa  3000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12359  GTCGTGATGCTTACCGAAGAACTGCTGCCTTATCCGCATAATCCGGCAAGCATTGAGCAA  12300

Qy    3001  gatcaggttgatttgatcgtcaaagttgaccgcgttggcgatgctgcaaaaatcggcgct  3060
            |||||||||||||||||||||||||||||||.||||||||||||||||||||||||||||
Db   12299  GATCAGGTTGATTTGATCGTCAAAGTTGACCGCGTTGGCGATGCTGCAAAAATCGGCGCT  12240

Qy    3061  ggcgcgacccgtatgaccactaacccgcgcgaactgcttattgcccgtagcgctgcggat  3120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12239  GGCGCGACCCGTATGACCACTAACCCGCGCGAACTGCTTATTGCCCGTAGCGCTGCGGAT  12180

Qy    3121  gtgattgtcaactctggctacttcaaagaaggtttctccatgcaaaccggcaccggcggc  3180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12179  GTGATTGTCAACTCTGGCTACTTCAAAGAAGGTTTCTCCATGCAAACCGGCACCGGCGGC  12120

Qy    3181  gcatcgctggcggtaacccgtttcctggaagacaaaatgcgtagccgcgatattcgcgcc  3240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12119  GCATCGCTGGCGGTAACCCGTTTCCTGGAAGACAAAATGCGTAGCCGCGATATTCGCGCC  12060

Qy    3241  gacttcgcccttggcggtattaccgcgacgatggttgacctgcacgaaaaaggtctgatc  3300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12059  GACTTCGCCCTTGGCGGTATTACCGCGACGATGGTTGACCTGCACGAAAAAGGTCTGATC  12000
```

```
Qy    3301  cgcaaactgctggatgtgcagagctttgacagccatgctgcgcaatcgctggcccgtaac  3360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11999  CGCAAACTGCTGGATGTGCAGAGCTTTGACAGCCATGCTGCGCAATCGCTGGCCCGTAAC  11940

Qy    3361  cccaatcacatcgaaatcagcgccaaccagtacgctaactggggttcgaaaggcgcatcg  3420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11939  CCCAATCACATCGAAATCAGCGCCAACCAGTACGCTAACTGGGGTTCGAAAGGCGCATCG  11880

Qy    3421  gttgatcgtctcgacgtggtggtactgagcgcgctggaaattgacacccagttcaacgtt  3480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11879  GTTGATCGTCTCGACGTGGTGGTACTGAGCGCGCTGGAAATTGACACCCAGTTCAACGTT  11820

Qy    3481  aacgtgctgaccggctctgacggcgtactgcgtggtgcttccggtggtcactgcgatacc  3540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11819  AACGTGCTGACCGGCTCTGACGGCGTACTGCGTGGTGCTTCCGGTGGTCACTGCGATACC  11760

Qy    3541  gcgattgcctctgcgctttccatcatcgtcgcgccgctggtacgcggtcgtattccgact  3600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11759  GCGATTGCCTCTGCGCTTTCCATCATCGTCGCGCCGCTGGTACGCGGTCGTATTCCGACT  11700

Qy    3601  ctggtggataacgtactgacctgcatcaccccaggctccagtgtcgatattctggtcaca  3660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11699  CTGGTGGATAACGTACTGACCTGCATCACCCCAGGCTCCAGTGTCGATATTCTGGTCACA  11640

Qy    3661  gaccacggtatcgcagttaacccggcacgtccggaactggcagaacgtctgcaggaagcg  3720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11639  GACCACGGTATCGCAGTTAACCCGGCACGTCCGGAACTGGCAGAACGTCTGCAGGAAGCG  11580

Qy    3721  ggcattaaagtggtttccattgagtggctgcgcgaacgtgcgcgtctgctgaccggtgaa  3780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11579  GGCATTAAAGTGGTTTCCATTGAGTGGCTGCGCGAACGTGCGCGTCTGCTGACCGGTGAA  11520

Qy    3781  ccacagccgattgaattcacagaccgcgtcgttgccgttgtgcgttaccgcgatggctcg  3840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11519  CCACAGCCGATTGAATTCACAGACCGCGTCGTTGCCGTTGTGCGTTACCGCGATGGCTCG  11460

Qy    3841  gtgatcgatgttgtgcatcaggtgaaggaataagccatgcacctgcttcctgaactcgcc  3900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11459  GTGATCGATGTTGTGCATCAGGTGAAGGAATAAGCCATGCACCTGCTTCCTGAACTCGCC  11400

Qy    3901  agccaccatgcggtatcaattcccgagctgctcgtcagccgggatgaaaggcaagcacgg  3960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11399  AGCCACCATGCGGTATCAATTCCCGAGCTGCTCGTCAGCCGGGATGAAAGGCAAGCACGG  11340

Qy    3961  caacacgtctggctcaagcgccatcctgttccactggtctcctttaccgtggttgcgcct  4020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11339  CAACACGTCTGGCTCAAGCGCCATCCTGTTCCACTGGTCTCCTTTACCGTGGTTGCGCCT  11280

Qy    4021  gggccgattaaagacagcgaggtcacacgccgaattttttaatcatggcgtgacagccttg  4080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11279  GGGCCGATTAAAGACAGCGAGGTCACACGCCGAATTTTTAATCATGGCGTGACAGCCTTG  11220

Qy    4081  cgtgccttagccgcaaaacagggctggcaaattcaggagcaggctgcactggtttccgcc  4140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11219  CGTGCCTTAGCCGCAAAACAGGGCTGGCAAATTCAGGAGCAGGCTGCACTGGTTTCCGCC  11160

Qy    4141  agcgggccggagggcatgttgagcattgccgccccggctcgcgacctcaagctcgccacc  4200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11159  AGCGGGCCGGAGGGCATGTTGAGCATTGCCGCCCCGGCTCGCGACCTCAAGCTCGCCACC  11100

Qy    4201  attgagcttgaacatagtcatcctctcgggcggttatgggatatcgatgtcctgacgccc  4260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11099  ATTGAGCTTGAACATAGTCATCCTCTCGGGCGGTTATGGGATATCGATGTCCTGACGCCC  11040

Qy    4261  gaaggcgaaattctctcccgccgcgactattcactgccgcctcgccgctgcctgttgtgc  4320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11039  GAAGGCGAAATTCTCTCCCGCCGCGACTATTCACTGCCGCCTCGCCGCTGCCTGTTGTGC  10980

Qy    4321  gaacaaagcgcagccgtctgcgcgcgtggaaaaacccatcaactgaccgatttactcaac  4380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10979  GAACAAAGCGCAGCCGTCTGCGCGCGTGGAAAAACCCATCAACTGACCGATTTACTCAAC  10920
```

```
Qy   4381 cgcatggaggcactgctgaacgatgtcgatgcctgcaacgtcaactaaaaccacaaagct 4440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10919 CGCATGGAGGCACTGCTGAACGATGTCGATGCCTGCAACGTCAACTAAAACCACAAAGCT 10860

Qy   4441 tgcgacgtcattaatcgatgagtacgccctgctgggctggcgcgccatgctgactgaagt 4500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10859 TGCGACGTCATTAATCGATGAGTACGCCCTGCTGGGCTGGCGCGCCATGCTGACTGAAGT 10800

Qy   4501 caatctgtcaccgaaaccaggcctcgtggatcgcattaactgcggtgcgcacaaagatat 4560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10799 CAATCTGTCACCGAAACCAGGCCTCGTGGATCGCATTAACTGCGGTGCGCACAAAGATAT 10740

Qy   4561 ggcgctggaagatttccaccgcagcgcgctggcgattcagggctggctaccccgtttcat 4620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10739 GGCGCTGGAAGATTTCCACCGCAGCGCGCTGGCGATTCAGGGCTGGCTACCCCGTTTCAT 10680

Qy   4621 tgaatttggtgcctgtagtgcggaaatggcaccagaagcggtactccacggattacgccc 4680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10679 TGAATTTGGTGCCTGTAGTGCGGAAATGGCACCAGAAGCGGTACTCCACGGATTACGCCC 10620

Qy   4681 aattggtatggcttgcgaaggtgatatgttccgcgccactgcgggcgtaaacacgcataa 4740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10619 AATTGGTATGGCTTGCGAAGGTGATATGTTCCGCGCCACTGCGGGCGTAAACACGCATAA 10560

Qy   4741 aggcagcattttttctttagggctgctatgtgcggcaattggccgtttgcttcaactcaa 4800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10559 AGGCAGCATTTTTTCTTTAGGGCTGCTATGTGCGGCAATTGGCCGTTTGCTTCAACTCAA 10500

Qy   4801 ccaaccggtaacgccaacaaccgtttgttctacggcggcaagtttctgccgtggcctgac 4860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10499 CCAACCGGTAACGCCAACAACCGTTTGTTCTACGGCGGCAAGTTTCTGCCGTGGCCTGAC 10440

Qy   4861 cgatcgcgaactgcgtaccaataattcacaactgacggcaggtcaacggttgtaccaaca 4920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10439 CGATCGCGAACTGCGTACCAATAATTCACAACTGACGGCAGGTCAACGGTTGTACCAACA 10380

Qy   4921 gcttggccttaccggcgcacgcggtgaagccgaagcgggttatccactggtgatcaatca 4980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10379 GCTTGGCCTTACCGGCGCACGCGGTGAAGCCGAAGCGGGTTATCCACTGGTGATCAATCA 10320

Qy   4981 cgccttgccgcattacctcactctgctggatcagggggttagatcctgaactggcattgct 5040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10319 CGCCTTGCCGCATTACCTCACTCTGCTGGATCAGGGGTTAGATCCTGAACTGGCATTGCT 10260

Qy   5041 cgataccttgctcctactgatggcgatcaacggcgataccaacgttgcatcgcgcggtgg 5100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10259 CGATACCTTGCTCCTACTGATGGCGATCAACGGCGATACCAACGTTGCATCGCGCGGTGG 10200

Qy   5101 cgagggggggcctgcgctggctacagcgcgaggcgcaaacattattgcaaaaaggggggcat 5160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10199 CGAGGGGGGCCTGCGCTGGCTACAGCGCGAGGCGCAAACATTATTGCAAAAAGGGGGCAT 10140

Qy   5161 tcgaaccccgccgatctcgattatctccggcagttcgacagggagtgtatcgaacgaaa 5220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10139 TCGAACCCCGCCGATCTCGATTATCTCCGGCAGTTCGACAGGGAGTGTATCGAACGAAA 10080

Qy   5221 tctcagtccaggcggcagtgctgacctactgatccttacctggttttttagcacagattta 5280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10079 TCTCAGTCCAGGCGGCAGTGCTGACCTACTGATCCTTACCTGGTTTTTAGCACAGATTTA 10020

Qy   5281 attatttaagcacttgataaatttggaaatattaattttcggagaacccgtatgtcttta 5340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10019 ATTATTTAAGCACTTGATAAATTTGGAAATATTAATTTTCGGAGAACCCGTATGTCTTTA 9960

Qy   5341 gcaaaagataatatatggaaactattggccccactggtggtgatgggtgtcatgtttctt 5400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9959 GCAAAAGATAATATATGGAAACTATTGGCCCCACTGGTGGTGATGGGTGTCATGTTTCTT 9900

Qy   5401 atccctgtccccgacggtatgccgccgcaggcatggcattacttcgctgtgtttgtggca 5460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9899 ATCCCTGTCCCCGACGGTATGCCGCCGCAGGCATGGCATTACTTCGCTGTGTTTGTGGCA 9840
```

```
Qy   5461 atgattgtcggcatgatcctcgag 5484
           ||||||||||||||||||||||||
Db    9839 ATGATTGTCGGCATGATCCTCGAG 9816
```